JSLHR

Research Article

Effects of Listener Age and Native Language Experience on Recognition of Accented and Unaccented English Words

Sandra Gordon-Salant,[a] Grace H. Yeni-Komshian,[a] Rebecca E. Bieber,[a] David A. Jara Ureta,[a] Maya S. Freund,[a] and Peter J. Fitzgibbons[a]

Purpose: Older native speakers of English have difficulty in understanding Spanish-accented English compared to younger native English speakers. However, it is unclear if this age effect would be observed among native speakers of Spanish. The current study investigates the effects of age and native language experience with Spanish on the ability to recognize words spoken in English by Spanish-accented and unaccented talkers.

Method: English monosyllabic words, recorded by native speakers of English and Spanish, were presented to 4 groups of listeners with normal hearing: younger native Spanish listeners (*n* = 15), older native Spanish listeners (*n* = 16), younger native English listeners (*n* = 15), and older native English listeners (*n* = 15). Speech recognition accuracy was assessed for the unaccented and accented words in both quiet and noise.

Results: In all conditions, the native English listeners performed better than the native Spanish listeners. More specifically, the native speakers of Spanish consistently recognized accented English less accurately than the native speakers of English, demonstrating no advantage of shared native language experience between nonnative listeners and accented talkers. Older listeners in the native Spanish language group also performed less accurately than their younger counterparts, for English words spoken by both unaccented and accented talkers. Finally, whereas listeners who were native speakers of English showed marked declines in recognition of Spanish-accented English relative to unaccented English, listeners who were native speakers of Spanish (both younger and older) showed less decline.

Conclusions: The general pattern of results suggests that both native language experience in a language other than English and age limit the ability to recognize Spanish-accented English. The implication of the overall findings is that older nonnative listeners will have considerable difficulty in understanding English, regardless of the talker's accent, in both clinical and everyday listening situations.

Among the many communication problems experienced by older listeners is difficulty in understanding accented speech. A number of investigations have demonstrated an age-related deficit in recognition of accented English that exceeds the age-related deficit observed for recognition of native English speech (Burda, Scherz, Hageman, & Edwards, 2003; Gordon-Salant, Yeni-Komshian, & Fitzgibbons, 2010a; Hargus Ferguson, Jongman, Sereno, & Keum, 2010). One theory offered to explain this age effect is that accented English is characterized by numerous changes, among them are the temporal characteristics of speech, which are not processed accurately by older people because of age-related auditory temporal processing deficits (Gordon-Salant et al., 2010a). That is, differences in expected timing of segmental and suprasegmental information in accented English relative to unaccented English, coupled with slowed auditory temporal processing by older listeners, combine to render accented English much less intelligible to older than to younger listeners. Many of the prior studies that evaluated perception of accented English by younger and older listeners presented speech produced by talkers whose native language was Spanish (Burda et al., 2003; Gordon-Salant et al., 2010a; Hargus Ferguson et al., 2010). Differences in temporal characteristics of Spanish-accented English, relative to unaccented English, have been reported, including variations in voice onset time (Flege & Eefting, 1988), vowel

[a]Department of Hearing and Speech Sciences, University of Maryland, College Park

Correspondence to Sandra Gordon-Salant: sgsalant@umd.edu

Editor-in-Chief: Frederick (Erick) Gallun
Editor: David Eddins

Received April 4, 2018
Revision received July 19, 2018
Accepted November 26, 2018
https://doi.org/10.1044/2018_JSLHR-H-ASCC7-18-0122
Publisher Note: This article is part of the Special Issue: Select Papers From the 7th Aging and Speech Communication Conference.

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

duration (Fox, Flege, & Munro, 1995), syllable stress (Adams & Munro, 1978), the amount of voicing during frication in voiced initial fricatives, and silent interval duration marking the fricative–affricate distinction (Gordon-Salant et al., 2010a). One investigation demonstrated that the types of recognition errors made by older listeners when recognizing Spanish-accented English on an open-set task were predominantly temporal in nature (Gordon-Salant et al., 2010a).

Listener experience has been shown to mitigate some of the age-related deficits observed for recognition of degraded speech. For example, Parbery-Clark, Strait, Anderson, Hittner, and Kraus (2011) reported that middle-age trained musicians performed significantly better than middle-age nonmusicians on word and sentence recognition tests presented in noise, as well as on an auditory temporal processing task (backward masking). The authors interpret the findings to indicate that musical training enhances the ability to understand speech in noise among an older cohort. Another investigation showed that listener experience with rapid speech eliminated the nearly ubiquitous finding of age-related deficits in recognition of time-compressed speech (Gordon-Salant & Friedman, 2011). In this prior study, older blind participants with normal hearing recognized time-compressed speech in quiet and noise significantly better than their sighted counterparts and not differently from younger sighted participants with normal hearing. Recognition scores for the most difficult time-compressed speech conditions were highly correlated with hours/week spent listening to accelerated speech (e.g., audiobooks or other text-to-voice materials) among the blind participants. These studies suggest that considerable focused listening experiences with certain types of acoustic and/or speech signals may minimize the expected difficulties that older people experience in recognizing degraded speech.

One type of experience that may affect the observed age-related difficulty in understanding accented English is being a native speaker of a language other than English and learning a second language later in life. In this article, the abbreviation "L1" is used to refer to individuals whose native language is English, and "L2" is used to refer to individuals whose second language is English. More specifically, the terms *L1* and *L2* are used throughout this article in reference to the native language background of the listeners. The terms *unaccented* and *accented* refer to the speech of talkers whose native language is English and Spanish, respectively. It may be posited that the age-related deficit in recognizing accented speech is reduced when the accented talker and L2 listener share the same native language. A number of studies have now shown that L2 listeners demonstrate a benefit for understanding accented English compared to L1 listeners (e.g., Bent & Bradlow, 2003; Hayes-Harb, Smith, Bent, & Bradlow, 2008; Weber, Broersma, & Aoyagi, 2011). Accented speakers of English often retain elements of the stress pattern of their native language and therefore produce sentences with a different tempo than unaccented speakers; in addition, accented speakers produce phonetic segments in English that may be dissimilar from those produced by unaccented speakers (Wenk, 1985). For example, for native Spanish speakers of English, distinct phonological differences, such as voice onset time, vowel duration, and so forth (relative to native English), permeate their accented productions of English words. Consequently, when L2 listeners hear Spanish-accented English, they hear many familiar components, because the accented speech contains elements that are characteristic of Spanish. Thus, an L2 listener's experience with the prosody and phonology of their native language may be expected to provide an advantage in recognizing foreign-accented English speech when there is a match between the native language of the speaker and listener. This phenomenon has been termed the *interlanguage speech intelligibility benefit* (ISIB; Bent & Bradlow, 2003), and the specific benefit characterized by better recognition by L2 listeners compared to L1 listeners for accented English speech has been called the *ISIB for listeners* (ISIB-L; Hayes-Harb et al., 2008). Participants in these previous studies, including both L1 listeners and L2 listeners, were young adult college students. Therefore, although the ISIB-L has been demonstrated for young adults, it is not clear if this benefit extends to older listeners as well.

A second form of ISIB is observed when L2 listeners recognize accented speech better than unaccented speech (Bent & Bradlow, 2003). This phenomenon has been termed the *ISIB for talkers* (ISIB-T; Hayes-Harb et al., 2008). It is another manifestation of the benefit afforded by foreign language experience and has been demonstrated by L2 adults listening to accented English sentences (Bent & Bradlow, 2003) and accented "English" monosyllabic nonwords (Xie & Fowler, 2013).

Numerous factors influence how well an individual acquires a second language, including age of arrival in the United States, age of second language acquisition, quality of second language input, and so forth (Flege, 2002). These factors, as well as spoken English proficiency by L2 listeners, have been shown to influence L2 listeners' recognition of English words produced by unaccented speakers (Imai, Walley, & Flege, 2005; Mayo, Florentine, & Buus, 1997; Shi, 2010). Although all of these factors are important, two were considered in the present investigation. The first was age of arrival in the United States. Participants were recruited who arrived after the age of 12 years, because studies have shown that those who arrive after this age retain their foreign accent (Yeni-Komshian, Flege, & Liu, 2000). The second factor is receptive English vocabulary, which was used as an approximate measure of English language proficiency (Chiang, 2018). Receptive English vocabulary is thought to play a major role in L2 listeners' recognition of spoken English (Calandruccio & Smiljanic, 2012) and was expected to be important in recognition of monosyllabic words that have no linguistic or contextual cues. The extent to which knowledge of English vocabulary contributes to recognition of unaccented and accented English speech by younger and older L2 listeners is unknown.

Other variables are known to contribute to the difficulty that L1 listeners experience when listening to

foreign-accented speech. In particular, the presence of noise has a significant impact on recognition of accented speech by L1 listeners (Munro, 1998) and may exacerbate age-related differences in recognition of accented speech (Gordon-Salant, Yeni-Komshian, & Fitzgibbons, 2010b). It is unclear if the age-related differences in recognizing accented speech in noise by older L1 listeners will be observed as well for older L2 listeners. Younger L2 listeners show lower recognition accuracy with decreases in the signal-to-noise ratio (SNR) for accented speech (Pinet & Iverson, 2010), but this effect has not been investigated for older L2 listeners. A significant decline in recognition of accented speech in noise by older L2 listeners may minimize the expected ISIB (ISIB-L) for these listeners.

Other important factors contributing to recognition of unaccented and accented English by older L1 listeners are their cognitive abilities and hearing sensitivity. Working memory capacity is a significant variable contributing to the difficulties of older listeners in understanding unaccented English (e.g., Akeroyd, 2008; Anderson, White-Schwoch, Parbery-Clark, & Kraus, 2013; Gordon-Salant & Cole, 2016; Schurman, Brungart, & Gordon-Salant, 2014) and accented speech (Bieber & Gordon-Salant, 2017; Janse & Adank, 2012), especially in noise. Other cognitive measures, such as processing speed and attention/inhibition, relate to perception of speech in difficult listening conditions, such as fast speech or in the presence of noise (Füllgrabe, Moore, & Stone, 2015; Gordon-Salant & Cole, 2016; Janse, 2009). The importance of cognitive abilities to recognition of unaccented and accented English speech by younger and older L2 listeners has not been examined previously. In the current investigation, the contribution of cognitive variables (working memory, speed of processing, and attention/response inhibition) to recognition of accented English by younger and older L1 and L2 listeners was examined. Other studies suggest that hearing sensitivity is the primary variable contributing significantly to recognition of accented speech in noise by L1 listeners (Gordon-Salant, Yeni-Komshian, Fitzgibbons, Cohen, & Waldroup, 2013). In the present investigation, participants with hearing sensitivity within normal limits were recruited in an effort to minimize the effect of hearing as a major variable.

The foregoing review covered the effects of numerous variables that have been shown to influence perception of unaccented and foreign-accented speech by L1 and L2 listeners. These variables include (a) the native language of the listener, (b) the native language of the talker, (c) the noise versus quiet testing condition, (d) the English language proficiency and vocabulary knowledge of the L2 listener, and (e) the cognitive abilities of the L1 listeners. Most of the prior work examining the impact of these variables was conducted with young adults, but it is known that older adults experience difficulty in understanding foreign-accented speech (e.g., Gordon-Salant et al., 2010a, 2010b). The general purpose of the current investigation was to examine the effect of native language experience of younger and older adults on perception of unaccented and foreign-accented speech. The study examined the effects of each of the variables listed above. The predictions for each of these variables are as follows:

1. Both younger and older L1 listeners recognize unaccented speech better than younger and older L2 listeners.
2. L2 listeners who share the same native language as the accented talkers recognize accented speech better than unaccented speech, and older L2 listeners will not show age-related decline for accented speech.
3. An interaction between listener language, talker language, and noise is predicted, in which older L2 listeners recognize accented speech better than older L1 listeners in quiet, but not in noise.
4. English proficiency of the L2 listeners, as measured by English vocabulary, is related to their recognition scores for the speech of unaccented and accented talkers.
5. Cognitive variables (working memory, speed of processing, and inhibition) contribute significantly to the variance in recognition of accented English by younger and older L1 and L2 listeners.

Method

Participants

Selection criteria for all participants were based on age (younger: 18–38 years; older: 60–80 years), audiometric indices consistent with normal hearing (based on audiometric thresholds and word recognition scores), native language (American English or Spanish), and a passing score on a cognitive screening tool. Data were collected from two groups of L2 listeners: young adults (ages 19–33 years, $M = 24.9$) and older adults (ages 61–80 years, $M = 69.4$). The first language of these L2 listeners was Spanish. L2 listeners were selected as individuals who were raised in a Spanish-speaking home and lived in a Spanish-speaking country until the age of at least 12 years. They subsequently arrived in the United States after that age (see Table 1). Data were also collected from two groups of L1 listeners: younger adults (ages 21–27 years, $M = 21.9$) and older adults (ages 67–76 years, $M = 69.3$). The L1 speakers of American English were raised in a monolingual, English-speaking home in the United States and did not report extensive recent travel abroad (i.e., semester study abroad in a non-English–speaking country during the past year). The L1 listeners reported essentially no listening, reading, writing, or speaking proficiency in Spanish. There were 16 listeners in the older native Spanish group and 15 listeners in each of the other three groups.

Audiometric criteria for recruitment into the study included normal audiometric thresholds, defined as pure-tone thresholds ≤ 25 dB HL (re: American National Standards Institute, 2010), from 250 to 4000 Hz, normal tympanograms, and acoustic reflex thresholds elicited at 0.5, 1, and 2 kHz in each ear. Word recognition scores (Northwestern

Table 1. Demographic information, self-ratings of English proficiency, and English vocabulary scores for the native Spanish listeners.

Listener	Age of arrival in the United States	Years of residence in the United States	Age (years)	Speaking proficiency (scale: 1–7)	Listening proficiency (scale: 1–7)	English vocabulary score
Younger						
SPA YNH1	25.0	1.0	26	7	7	152
SPA YNH2	14.0	5.0	19	6	7	165
SPA YNH4	14.0	11.0	26	7	7	159
SPA YNH5	27.0	6.0	33	6	6	158
SPA YNH6	20.0	1.0	21	6	6	166
SPA YNH7	24.0	1.0	25	6	6	139
SPA YNH8	25.0	5.0	30	5	6	152
SPA YNH9	13.0	16.0	29	6	6	161
SPA YNH10	22.0	1.0	23	6	6	168
SPA YNH11	15.0	7.0	22	5	6	156
SPA YNH12	22.0	0.7	22	5	6	175
SPA YNH13	23.0	0.7	23	6	6	147
SPA YNH15	22.0	0.7	22	6	6	145
SPA YNH16	17.0	12.0	30	6	6	145
SPA YNH17	21.0	0.5	21	6	6	147
Older						
SPA ONH1	29.0	49.0	78	7	6	165
SPA ONH2	20.0	51.0	71	3	4	138
SPA ONH3	49.0	19.0	68	4	4	127
SPA ONH4	69.0	1.0	70	4	4	104
SPA ONH5	34.0	36.0	70	7	7	157
SPA ONH7	25.0	48.0	73	5	5	84
SPA ONH8	21.0	49.0	70	6	7	132
SPA ONH9	20	44	64	4	6	90
SPA ONH10	52.0	17.0	69	3	3	92
SPA ONH11	50.0	30.0	80	3	3	92
SPA ONH12	32.0	30.0	61	3	3	128
SPA ONH13	33.0	30.0	63	*	*	95
SPA ONH14	24.0	42.0	67	*	*	135
SPA ONH15	53.0	7.0	60	4	4	93
SPA ONH16	51.0	16.0	67	3	5	89
SPA ONH17	45.0	17.0	61	3	3	163

Note. Self-ratings are based on a scale of 1–7, with 1 = *very poor* and 7 = *nativelike*. SPA = native Spanish listeners; YNH = younger listeners with normal hearing; ONH = older listeners with normal hearing.

*Indicates no rating provided.

University Auditory Test No. 6; Tillman & Carhart, 1966) were at least 80% at a presentation level of 75 dB HL to ensure that listeners did not have atypical speech recognition scores. Average word recognition scores exceeded 89% for all groups. Audiograms of the four listener groups are shown in Figure 1.

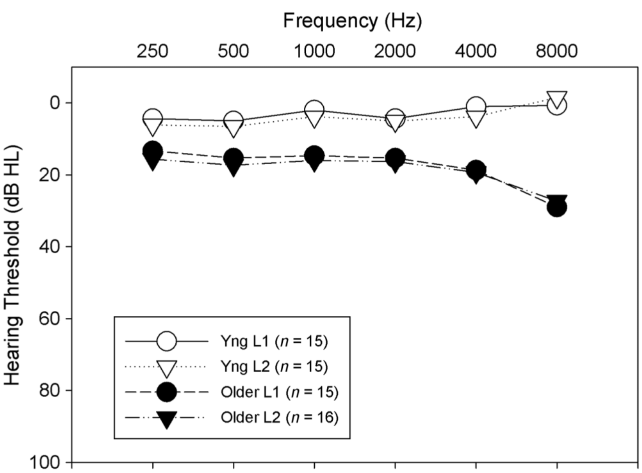
All listeners were required to pass a screening test of general cognitive awareness. The test administered to the L1 listeners was the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975), and the test administered to the L2 listeners was the MMSE-2, which is the Spanish version of the MMSE. A passing score of 24 or higher was required for recruitment into the study, indicating normal cognitive awareness.

Listeners of all four groups were required to have at least a high school diploma (either from the United States or the participant's home country). Additional criteria for selection of the L2 listeners were residence in the United States for at least 6 months and age of arrival in the United States of 12 years or older. Proficiency in English was not a selection criterion, although extensive information about native and second language histories was collected from all participants. This included self-ratings of reading, writing, speaking, and listening proficiency in the second language. Table 1 presents demographic information about the L2 listeners, including the age of arrival and years of residence in the United States, as well as self-ratings of speaking and listening proficiency in English (scale of 1–7 with 1 = *very poor* and 7 = *nativelike*). A repeated-measures analysis of variance of the proficiency ratings shown in Table 1 indicated that younger L2 listeners rated their speaking and listening proficiency higher than older L2 listeners, $F(1, 27) = 19.21$, $p < .001$, and all L2 listeners rated their listening proficiency higher than their speaking proficiency, $F(1, 27) = 6.26$, $p < .05$.

Stimuli

The stimuli used in these experiments were monosyllabic words as described in detail in an earlier study (Gordon-Salant et al., 2010a). Briefly, there were 160 consonant–vowel–consonant English words, assigned to

Figure 1. Pure-tone audiometric thresholds in dB HL (re: American National Standards Institute, 2010) for the four listener groups.



equivalent lists of 40 words each that contained numerous consonant and vowel phonemes that are often pronounced differently by native speakers of Spanish and native speakers of English. Each word was one of a contrasting pair, such that a misperception of one phoneme could result in perception of another English word (e.g., "safe" and "save"). These contrasting word pairs appeared on different stimulus lists. The 160 words were recorded by an unaccented speaker of English and two accented speakers of English (native language = Spanish), who were young male adults (ages 19–25 years). For this study, only the recordings of the more accented speaker were used. The mean rating of his accentedness by 10 young normal hearing L1 listeners was 3.62, on a 5-point scale, with 1 = *no accent* and 5 = *a severe accent*, indicating he was perceived as a moderately accented speaker. The mean rating of accentedness for the unaccented speaker was 1.04, as expected. The monosyllabic word stimuli were equated for root-mean-square level, and a calibration tone was created to be equal in root-mean-square level to that of the word stimuli. Four separate lists of the monosyllabic words (40 words each) were created for each talker.

A background noise of six-talker male babble was used for the noise conditions. Details of the creation of this babble are described elsewhere (Gordon-Salant et al., 2013). The babble was continuous, whereas the word stimuli (including a carrier phrase, "Number *x*" where *x* corresponded to the number of the stimulus on the list) were inserted digitally on the speech channel with a 4-s interstimulus interval. The speech stimuli and background noise were burned to separate channels of compact disks, preceded by calibration tones for each channel (speech and noise).

Cognitive and Vocabulary Measures

Five cognitive measures were administered to listeners to determine if variation in cognitive performance was related to performance on the speech recognition measures. Two cognitive measures assessed working memory: the Forward and Backward Digit Span from the Wechsler Adult Intelligence Scale–Third Edition (Wechsler, 1997); these measures were selected because they did not require extensive knowledge of English (Ramkissoon, Proctor, Lansing, & Bilger, 2002). Two cognitive measures assessed speed of processing: the Symbol Search and the Digit Symbol Coding subtests of the Wechsler Adult Intelligence Scale–Third Edition (Wechsler, 1997). The fifth cognitive measure was the Flanker Test (Eriksen & Eriksen, 1974), a test of response inhibition. The Flanker Test was downloaded from the National Institutes of Health Cognitive Toolbox (Weintraub et al., 2013) for this study. In addition to the cognitive measures, participants were assessed on their knowledge of English vocabulary. The test used for this purpose was the Receptive One-Word Picture Vocabulary Test–Fourth Edition (ROWPVT-4; Martin & Brownell, 2011). The ROWPVT-4 presents monosyllabic and multisyllabic words to listeners that are nouns, verbs, and adjectives. Listeners select the word heard from a closed set of four choices presented pictorially. For these adult listeners, testing began with the vocabulary items recommended for the oldest age group (age = 14;0 [years;months] and older) and a baseline score of eight consecutive correct items was required before testing continued. Testing terminated when the listener made six incorrect responses out of the last eight words presented. All listeners met the baseline requirement.

Procedure

Prior to participating in the experimental conditions, listeners completed several tests, including the audiometric evaluation, acoustic immittance screening, the Language History Questionnaire, and the MMSE. If they met subject selection criteria, they completed the cognitive measures and the ROWPVT-4.

The four experimental conditions consisted of monosyllabic words recorded by the unaccented and accented talkers presented in quiet and noise (i.e., multitalker babble) at a + 5 dB SNR. The + 5 dB SNR was chosen following pilot testing as the SNR that yielded, on average, a 20% reduction in score from the quiet to the noise condition (i.e., 80% correct) for five young L1 listeners, when presented with unaccented speech. The signal level for all conditions was 85 dB SPL. The speech stimuli and background babble were routed in separate channels from a Tascam CD player (model CDRW-402) to a Crown D-75 amplifier, Hewlett-Packard 350 D attenuator, and a Colbourn audio-mixer amplifier (model S82-24). The mixed signal was then delivered to one ear of the listener (right ear or better ear), consistent with previous investigations of perception of accented speech, to facilitate comparison of data across studies (e.g., Gordon-Salant et al., 2010a). The transducer was an Etymotic ER-3A insert earphone. Listeners were seated in a sound-attenuating booth during the experimental conditions. Their task was to repeat and write down the word they heard. Scoring was based on the listener's verbal response.

The written response was available if the listener had a heavy accent and the verbal response was unclear. In addition, the listener's verbal response was recorded for later verification if the written response was illegible. Nevertheless, the need to rely on the written and/or recorded response was rare.

The presentation of the speech conditions was blocked by quiet versus noise condition, and the block order was alternated between participants. Within a block, the order of presentation of the lists recorded separately by the two talkers (unaccented, accented) was randomized across participants. Finally, the order of stimulus lists (1–4) was randomized using a Latin square design. Thus, each participant received a unique order of noise condition, talker accent, and stimulus list to eliminate any systematic learning effects.

Results

Effects of Talker Native Language, Listener Native Language, Age, and Noise

Monosyllabic word recognition scores of the four listener groups in quiet and noise conditions for the unaccented and accented talkers are shown in Figure 2 and Table 2. Effects of native language experience and listener age are apparent, with L1 listeners performing better than L2 listeners, and younger listeners performing better than older listeners. However, the pattern of group performance also varies with the native language of the talker. A statistical model was fit to the word recognition data using a generalized linear mixed-effects regression analysis (glmer) in the lme4 package with R studio software (Bates, Mäechler, Bolker, & Walker, 2015), following the general principle to include maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013). The dependent variable was the trial-by-trial binomial word recognition score (correct or incorrect). The initial base model testing included all fixed effects in the design (listener native language, listener age, talker native language, quiet/noise condition, and interactions between them) and selected random effects (intercepts and slopes for participants in the quiet/noise conditions and in the talker language conditions). Model testing proceeded with iteratively removing fixed and random effects that did not contribute significantly to the model. The test statistic is expressed as a *z* value or *z* ratio value (parameter estimate divided by the standard error) in multilevel modeling (i.e., glmer; Finch, Bolin, & Kelley, 2014). The interpretation of the *z* values is comparable to that for the *t* statistic.

Figure 2. Mean recognition scores for monosyllabic words spoken by the unaccented talker and the Spanish-accented talker presented in quiet and noise, by the four listener groups. Error bars represent 1 *SE*.

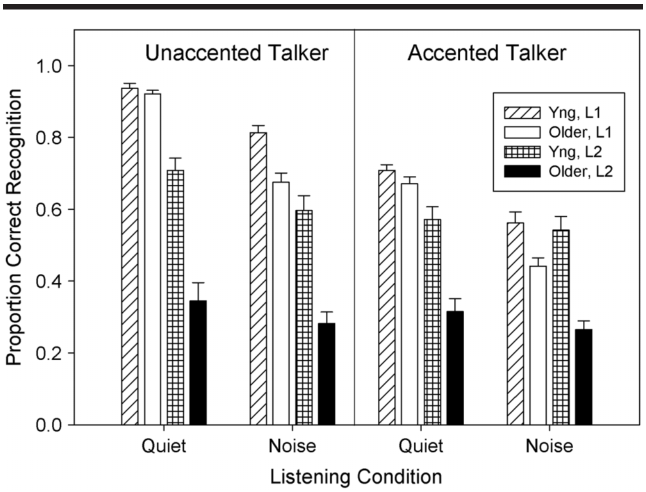


Table 2. Mean recognition performance (and standard deviations) of the four listener groups in four listening conditions.

Listener group	Native English talker				Native Spanish talker			
	Quiet		Noise		Quiet		Noise	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Younger, L1	.94	.05	.81	.08	.71	.06	.56	.12
Older, L1	.90	.04	.69	.03	.68	.02	.49	.12
Younger, L2	.70	.13	.60	.16	.57	.14	.54	.15
Older, L2	.37	.18	.30	.11	.34	.13	.28	.08

Note. L1 refers to listeners whose native language is English, and L2 refers to listeners whose native language is Spanish.

The best-fitting model derived from these fixed and random effects included the random coefficient for noise per subject, and slope estimates for the fixed effects showed that L2 listeners performed less accurately than L1 listeners ($\beta = -2.04$, $z = -12.44$, $p < .0001$) and that younger listeners performed better than older listeners ($\beta = 0.60$, $z = 3.68$, $p < .0001$). In addition, all listeners obtained lower recognition scores for accented talkers than for unaccented talkers ($\beta = -0.97$, $z = -10.71$, $p < .0001$) and performed better in the quiet conditions than in the noise conditions ($\beta = 1.21$, $z = 10.96$, $p < .0001$). However, significant two-way interactions were also revealed in the model: Listener Language × Talker Language ($\beta = 1.14$, $z = 11.84$, $p < .0001$), Age Group × Talker Language ($\beta = -0.39$, $z = -4.2$, $p < .0001$), Listener Language × Noise Condition ($\beta = -0.68$, $z = -5.54$, $p < .0001$), Talker Language × Noise Condition ($\beta = -0.36$, $z = -3.78$, $p < .0001$), and Listener Language × Age Group ($\beta = 0.83$, $z = 3.81$, $p < .0001$). None of the higher order interactions were significant.

Post hoc analyses were conducted to evaluate the two-way interactions further using the function lsmeans in R. As noted above, the effect of talker native language interacted with listener native language and age. Post hoc analyses using the Tukey method to adjust the *p* value for comparing a family of estimates examined these interactions. The post hoc analyses of the Listener Language × Talker Language interaction, with results collapsed over noise conditions and age groups, indicated that there was

a larger listener language effect for the unaccented talker ($z = 17.58$, $p < .0001$) than for the accented talker ($z = 7.23$, $p < .0001$). The mean group scores, shown in Table 2, indicate that the ISIB-L was not observed for these monosyllabic word stimuli. That is, the L2 listeners did not show better recognition of the accented talker than the L1 listeners, as was predicted by the ISIB-L hypothesis. The interaction between listener language and talker language was also driven by a larger effect of talker language for the L1 listeners ($z = 16.17$, $p < .0001$) than for the L2 listeners ($z = 3.70$, $p < .01$). Post hoc analyses of the Age × Talker Language interaction showed that the magnitude of the age effect was smaller for the accented talker than for the unaccented talker, as reflected by the comparison of absolute z values (9.84 for unaccented talker vs. 5.62 for the accented talker, $p < .0001$, both comparisons).

There were 2 two-way interactions involving noise. The significant interaction between listener language and noise was analyzed with the Tukey method for multiple comparisons by averaging results over the levels of age group and talker native language. Post hoc testing showed that the decline in performance from quiet to noise was larger for L1 listeners than L2 listeners based on absolute z values (L1 listeners: $z = |-11.34|$, $p < .0001$; L2 listeners: $z = |-4.2|$, $p < .001$). This finding suggests that L2 listeners were less affected by noise than L1 listeners when listening to English speech, regardless of the talker's native language. As for the Talker Language × Noise interaction, post hoc testing indicates that the source of the interaction was a smaller difference in performance in the quiet versus noise conditions for the accented talker ($z = -4.2$, $p < .001$) compared to the unaccented talker ($z = -11.34$, $p < .0001$).

The final interaction was between listener language and age group. This interaction was analyzed by collapsing the results over the levels of noise condition and talker language, again using the Tukey method for multiple comparisons. Post hoc testing showed that younger L2 listeners showed higher accuracy scores than older L2 listeners ($z = -8.64$, $p < 0001$), but younger and older L1 listeners did not differ in their accuracy scores ($z = -2.27$, $p > .05$). In addition, older L1 listeners obtained higher scores than older L2 listeners ($z = 13.07$, $p < .0001$), and younger L1 listeners obtained higher scores than younger L2 listeners ($z = 5.36$, $p < .0001$). Thus, the source of the interaction is a significant age effect for L2 listeners, but not L1 listeners.

Predictors of Recognition Performance for the Four Listener Groups

The final analysis evaluated the extent to which predictor variables of hearing sensitivity, cognition, and English vocabulary improved the model describing the significance of variables contributing to word recognition performance. Model testing with general linear mixed-effects regression analysis proceeded from the initial model described above to subsequently include, in separate iterations, each of the additional predictor variables collected from the listeners. These additional predictor variables (all continuous variables) included measures of working memory (Forward Digit Span, Backward Digit Span), speed of processing (Digit Symbol Coding, Symbol Search), attention/inhibition (Flanker), English vocabulary (ROWPVT-4), and pure-tone hearing thresholds (quantified as pure-tone average of thresholds at .5, 1, and 2 kHz [the PTA] and as high-frequency pure-tone average of thresholds at 1, 2, and 4 kHz [the HF-PTA]). Each of the scores on these measures was converted to a z score prior to entering each variable sequentially into the model. The only predictor variable that improved the model was the English vocabulary score ($\beta = 0.36$, $z = 4.31$, $p < .0001$), as verified by an analysis of variance ($\chi^2 = 16.37$, $p < .001$) that compared the model without the inclusion of the English vocabulary score to the model with the English vocabulary score. To estimate how much the speech recognition score increases with an increase in English vocabulary score, the "divide by four rule" was applied (Gelman & Hill, 2007). This rule of convenience is used to determine the upper bound of the predictive difference near the midpoint of a logistic curve, which is steepest at the center where probabilities approximate .5. The calculation is accomplished by dividing the β coefficient by four. Applying this rule to the coefficient for the English vocabulary z score (0.36) suggests that the maximum increase in the predicted speech recognition score is 9% with each unit vocabulary z score increase (i.e., each standard deviation above the mean vocabulary z score of 0.5).

Discussion

This investigation of the effects of native language experience, talker accent, and listener age sought to determine if native language experience in the same language as an accented talker is beneficial for recognizing foreign-accented speech and if this benefit reduces age effects observed for understanding accented speech. It was predicted that younger and older L2 listeners recognize Spanish-accented English better than L1 listeners of comparable ages, because of their familiarity with the characteristics of Spanish that are presumably retained during production of Spanish-accented English. It was also predicted that L2 listeners would not show age-related differences in understanding accented speech, as has been shown previously for L1 listeners, because of the expected advantage of a shared native language with the talker. The findings generally do not support these predictions. Rather, L2 listeners did not recognize accented speech better than L1 listeners, and older L2 listeners performed more poorly than younger L2 listeners.

Native Language Experience of the Listener

The first objective of this investigation was to determine if a listener's native language experience with a foreign language (Spanish) provides an advantage for recognition of accented English spoken by a talker who

shares the same native language as the listener. The results showed that L2 listeners did not recognize accented words better than L1 listeners, but rather L2 listeners consistently showed poorer recognition of accented words compared to L1 listeners. These findings contradict the theory that listeners who share a native language with an accented talker will derive an ISIB (ISIB-L; Bent & Bradlow, 2003; Hayes-Harb et al., 2008).

There are several possible reasons for differing results across studies. One reason why the findings of the current study were not consistent with previous studies may be related to the English language proficiency of the talkers. In a previous investigation, the ISIB-L was observed for the speech of accented talkers who had a high level of English speaking proficiency, but not for accented talkers with a low level of English speaking proficiency, as determined by ratings of production speech intelligibility scores (Bent & Bradlow, 2003). Strength of foreign accent was not specifically examined in this study. The current investigation included a single accented speaker. He was rated as having a moderate accent by a group of L1 listeners. However, a different group of L1 listeners rated his speech intelligibility for sentences to be approximately 84%. This intelligibility rating is comparable to that of the highly proficient accented talkers in the Bent and Bradlow (2003) study. It is possible that the strength of foreign accent reflects a more important dimension of talker proficiency for perception of single word stimuli such as those used in the current study. Additional investigation of the ISIB-L as related to different indices of talker proficiency for different speech materials will likely clarify this issue.

A second reason for contrasting findings in this study compared to others may be differences in the speech materials and listening task. The current study presented monosyllabic words in an open-set task to listeners as a first attempt to investigate the effects of language experience and age on recognition of accented speech. The choice of using monosyllabic words from dense lexical neighborhoods coupled with an open-set task in the current study forced listeners to rely on the acoustic–phonetic cues of speech without lexical or contextual support, including the stress patterns of the native language that often transfer to sentence-length test materials. Understanding monosyllabic words from dense lexical neighborhoods is challenging for L2 listeners, especially when coupled with an open-set task (Clopper, Pisoni, & Tierney, 2006). The present findings indicate that L2 listeners do not obtain significant benefit from listening to an accented talker who shares the same native language for these materials and this type of task. Prior studies presented sentences with an open-set recognition task (Bent & Bradlow, 2003), minimal pairs of monosyllabic words with a two-alternative forced-choice task (Hayes-Harb et al., 2008), or monosyllabic nonwords (Imai et al., 2005; Xie & Fowler, 2013). Thus, in comparison to previous studies, the design of the current study may have presented test conditions in which an ISIB-L is less likely to be observed. The current findings suggest that the ISIB-L is not a strong and consistent phenomenon but, rather, is specific to the stimulus characteristics and response task.

Talker Native Language

A second goal of this investigation was to determine if L2 listeners exhibit greater recognition accuracy for accented speech compared to unaccented speech (ISIB-T). This was not observed in the present investigation. Both L1 and L2 listeners recognized unaccented words better than accented words. Thus, there was no benefit of native experience with Spanish for listening to English monosyllabic words produced by a Spanish-accented talker; in other words, predictions based on the ISIB-T were not supported. This observation is in contrast to the findings of Bent and Bradlow (2003). Several differences in the research design between the two studies may have contributed to the discrepant results. One distinction is that Bent and Bradlow recruited accented talkers at two levels of English speaking proficiency (high and low). The ISIB-T effect reported in that study was observed for the highly proficient accented speaker and not for the less proficient speaker. It is unclear if the language proficiency of the single accented talker in the current study would be rated as high or low, based on differing criteria as noted earlier. If his English speaking proficiency was similar to that of Bent and Bradlow's less proficient talker, this may explain the absence of the ISIB-T reported here. In addition, the speech materials differed between the two studies (single words in the current study and sentences in the Bent and Bradlow study), and prior studies suggest that the observation of an ISIB-T may be dependent on the type of speech material. For example, Hayes-Harb et al. (2008) presented unaccented and accented English monosyllabic words to L1 and L2 listeners (native language of the accented talkers and L2 listeners: Mandarin Chinese) and did not observe an ISIB-T.

In addition to type of speech material, the English proficiency of the "listeners" also influences the observation of an ISIB-T. A study by Xie and Fowler (2013) presented unaccented and Mandarin-accented "English" monosyllabic nonwords to L1 and L2 listeners (native language of L2 listeners: Mandarin Chinese) and reported an ISIB-T for L2 listeners with low English proficiency who resided in Beijing, but not for low English proficiency L2 listeners who resided in the United States. L2 listeners with high English proficiency residing in either country did not demonstrate an ISIB-T. The English language proficiency of the L2 listeners in the current study was variable, as this was not a recruitment criterion. In summary, the current findings reinforce and extend prior results indicating that the ISIB-T is a relatively weak phenomenon, which is dependent on the speech materials, and specific speaker and listener characteristics. The present findings indicate that the ISIB-T is not seen for monosyllabic words nor for L2 listeners with a range of language proficiency.

Even though the L2 listeners recognized accented speech less accurately than unaccented speech, the decline

in performance with the accented talker was not as large for L2 listeners as for L1 listeners. All L1 listeners showed a substantial decrease in recognition of accented speech compared to unaccented speech, as shown in previous studies (Gordon-Salant et al., 2010a, 2010b). However, the magnitude of the decrease in performance scores for the accented talker compared to the unaccented talker was much reduced for the L2 listeners. Examination of the marginal means presented in Table 3 shows that the decline in scores for the accented talker compared to the unaccented talker was 6% for the L2 listeners and 23% for the L1 listeners. This observation may be considered as an indirect and weak support for the ISIB-T; in other words, the "penalty" of listening to accented speech (i.e., decline in recognition of accented vs. unaccented speech) appears to be less for L2 listeners than for L1 listeners. The findings may be driven, in part, by the relatively high recognition scores for unaccented speech by the L1 listeners and the difficulty that they experience in listening to accented speech, resulting in a large decrease in performance with accented speech. In contrast, the L2 listeners appear to have difficulty in understanding English words, regardless of talker accent, resulting in a more modest difference in recognition scores for unaccented and accented speech. That is, listening to English speech is difficult and effortful for the L2 listeners and may render them less sensitive to further changes in the speech signal with accent. Nonetheless, the current findings clearly show that all listeners, both L1 and L2, have more difficulty in understanding accented speech than unaccented speech, which is not consistent with the claims of the ISIB-T.

Listener Age

Another major objective of this investigation was to determine if the ISIB-L minimized age-related differences in L2 listeners compared to L1 listeners. The expectation was that L2 listeners would benefit from experience with the same native language as the accented talker (i.e., Spanish) for understanding accented speech, and as a result, this benefit would reduce the age effects often observed for L1 listeners during recognition tasks with accented speech. This was not observed. Rather, younger listeners performed better than older listeners for both the unaccented and accented talkers, but younger listeners were more impacted by talker accent than older listeners. That is, younger listeners (both L1 and L2) showed greater declines in performance than older listeners (L1 and L2) for recognition of words spoken by the accented talker compared to the unaccented talker. The observation that older listeners' performance did not decline as much as younger listeners' performance in the accented speech condition may in part reflect the fairly stable (and low) performance across talker conditions of the older L2 listeners. In fact, the low performance level of the older L2 listeners did not show substantial variation across all speech conditions. A key finding in this investigation is a significant age effect for L2 listeners, but not for L1 listeners. This finding refutes the prediction that native language experience in a language other than English (i.e., as observed in the L2 listeners) would mitigate age-related differences in listening to accented speech observed previously for L1 listeners (Gordon-Salant et al., 2010a, 2010b). L1 listeners did not show the expected age effect for accented speech, either in quiet or in noise. This may be related to selection of older listeners with normal hearing and the use of a relatively high SNR in the noise condition. These methodological choices were made to minimize potential floor effects that might have been observed with older L2 listeners presented with a lower SNR. The absence of an age effect for L1 listeners with normal hearing in the current study suggests that some of the age-related performance deficits reported in prior investigations may be related more to loss of hearing sensitivity than to age per sé.

Although many previous reports examined the effect of age on recognition of accented English (Burda et al., 2003; Gordon-Salant et al., 2010a; Hargus Ferguson et al., 2010) and the effect of native language on recognition of native and accented English words (Mayo et al., 1997; Nakamura & Gordon-Salant, 2011; Shi, 2010), there are no prior reports that evaluated the effects of both native language experience and listener age on word recognition in the same study. Taken together, the current findings strongly suggest that age is an important listener attribute to consider when assessing the performance of L2 listeners, because older L2 listeners have even greater problems perceiving accented speech than younger L2 listeners. In addition, native language experience is an important factor to consider when assessing the performance of older listeners, as older L2 listeners experience more difficulty when listening to unaccented and accented speech than older L1 listeners. Even though all listeners in this study had normal hearing, the older L2 listeners performed considerably worse than the younger L2 listeners and considerably worse than the older L1 listeners, suggesting that the effects of age and native language are somewhat additive when listening to English words. Previous studies have not examined the performance of older L2 listeners on English word recognition tasks and, thus, failed to demonstrate the significant difficulties experienced by older L2 listeners in understanding unaccented English words presented in either quiet or noise.

Table 3. Marginal mean recognition scores (proportions and standard deviations) for L1 and L2 listeners responding to unaccented and accented English.

Listener group	Talker native language			
	Unaccented English		Accented English	
	M	*SD*	*M*	*SD*
L1	.84	.07	.61	.07
L2	.49	.21	.43	.17

Note. L1 refers to listeners whose native language is English, and L2 refers to listeners whose native language is Spanish.

Quiet Versus Noise

The investigation also examined if noise magnifies age-related differences in word recognition performance for L2 listeners. This finding was not observed, indicating that younger versus older L2 listeners were not differentially affected by noise compared to younger versus older L1 listeners. Rather, the effect of listener language varied with noise condition. This interaction appears to be driven by a very large performance difference between quiet and noise conditions for L1 listeners, which was reduced for L2 listeners. As noted above, it appears that, for L2 listeners, understanding English is difficult and effortful, regardless of noise condition, which may explain the limited shift in performance between quiet and noise conditions. However, for L1 listeners, the addition of noise has a large impact, consistent with many previous studies (e.g., Dubno, Dirks, & Morgan, 1984; Stuart & Phillips, 1996). An alternative way to view this interaction is that the effect of listener language was significant not only in noise (as expected based on previous investigations; e.g., Mayo et al., 1997; von Hapsburg, Champlin, & Shetty, 2004) but also in quiet. Most prior studies report that L2 listeners perform similarly to L1 listeners in quiet (e.g., Tabri, Abou Chacra, & Pring, 2011; von Hapsburg et al., 2004). The current finding that the L2 listeners performed more poorly than the L1 listeners in quiet may be associated with the use of monosyllabic word stimuli, rather than sentences, which were used in many previous studies (Broersma & Scharenborg, 2010; Mayo et al., 1997; Tabri et al., 2011; von Hapsburg et al., 2004). Sentence materials provide contextual cues for listeners, and L2 listeners have been shown to take advantage of contextual cues (Bradlow & Alexander, 2007). Thus, the use of monosyllabic word stimuli in the current study did not permit L2 listeners to rely on context, which may have aided their performance in other studies. In addition, the monosyllabic words selected for the current investigation generally came from dense lexical neighborhoods, which may have compounded the difficulty of the word recognition task for the L2 listeners even in quiet (Luce & Pisoni, 1998). The findings suggest that, in addition to the difficulties in understanding speech in noise conditions reported previously, L2 listeners may have considerable difficulty in understanding speech in quiet for monosyllabic words devoid of contextual cues.

The effect of noise, relative to quiet, was smaller for the accented talker than the unaccented talker. This suggests that listening to a foreign-accented talker in quiet imposes challenges for listeners and that the addition of noise, although significant, is not as detrimental as for unaccented talkers. Nevertheless, combining the findings that all listeners show poorer scores for accented than unaccented speech and poorer scores in noise than in quiet indicates that listening to accented talkers in noise is a markedly difficult listening condition, regardless of age and native language background. This observation is reinforced by the accuracy scores shown in Table 2 and Figure 2, in which mean recognition scores for the accented talker in noise are quite low (< 60% correct for young listeners and < 50% correct for older listeners).

Cognition and Knowledge of English Vocabulary in Relation to Recognition of Accented Speech

A final goal of this investigation was to determine the extent to which performance on tests of English vocabulary and selected cognitive abilities predicts recognition of accented English and whether or not significant predictive measures for accented English differ from those that predict recognition of unaccented English. In the current investigation, the linear mixed-effects regression model evaluated the significance of each of the cognitive performance measures of working memory, speed of processing, and attention/inhibition, in addition to hearing sensitivity and English vocabulary score, in improving the model. Among these measures, the single variable that significantly improved the baseline model was English vocabulary score (ROWPVT-4). This indicates that when the data of both L1 and L2 listeners are evaluated, knowledge of English vocabulary is a significant predictor of recognition performance for both unaccented and accented speech. Previous research suggests that vocabulary knowledge is an important variable contributing to recognition of accented speech by L1 listeners (Banks, Gowen, Munro, & Adank, 2015). However, the current findings appear to be largely determined by the English vocabulary scores of the L2 listeners, whose demographic information (shown in Table 1) indicates that these listeners exhibited a wide range of English vocabulary scores. The large variability in English vocabulary scores was especially apparent among the older L2 listeners, despite their residence in the United States for many more years than the younger L2 listeners. Overall, these findings demonstrate that for L2 listeners, knowledge of the lexicon is the most important predictor variable (among those evaluated) of recognition of English words. Knowledge of the English lexicon is important for word recognition by L2 adults even when those words occur with high frequency in the language and when the L2 listeners have resided in the United States for numerous years, which suggests a long period of immersion in the English language.

The cognitive variables of working memory capacity, processing speed, and attention/inhibition were not significant predictors of performance in this study, contrary to many previous studies (Füllgrabe et al., 2015; Gordon-Salant & Cole, 2016; Gordon-Salant et al., 2013; Schurman et al., 2014; Vaughan, Storzbach, & Furukawa, 2006). One possible reason is that the speech materials in most previous studies were unaccented sentences that were either speeded or presented in noise (e.g., Füllgrabe et al., 2015; Schurman et al., 2014; Vaughan et al., 2006), which may have taxed older listeners' working memory capacity or processing speed to a greater extent than the monosyllabic words used in the current investigation. Another possible reason is that the working memory test chosen for this study, Digit Span, is considered a simple verbal storage task, and adults

typically perform better on this test than on complex verbal working memory tests that involve both a simple span task plus a secondary processing task (Hale et al., 2011). Thus, the choice of working memory test may not have been sufficiently sensitive to reveal the contribution of variations in working memory to recognition of accented speech. However, at least one prior study examined predictor variables for recognition of accented speech and reported that cognitive performance did not significantly predict recognition of accented speech by L1 listeners (Gordon-Salant et al., 2013). Rather, the only significant predictor variable of recognition of accented speech by younger and older L1 listeners in that study was HF-PTA (Gordon-Salant et al., 2013). In the current study, neither PTA nor HF-PTA was a significant predictor of performance, as was expected given that the participants all had normal hearing sensitivity.

The current findings have major clinical implications, in that foreign-born clinical patients may perform exceptionally poorly on speech recognition tests administered in quiet and in noise in the Audiology Clinic. With the aging of the population and accompanying age-related hearing loss, clinicians may expect an increasing number of older nonnative speakers of English to seek audiologic services. Even though some of these patients may have resided in the United States for decades, their performance on standard speech recognition tests may still be affected by their more limited knowledge of the English lexicon compared to older native English listeners. Additional research on the multiple linguistic and situational factors that may influence speech recognition outcomes among older L2 listeners with both normal hearing and with hearing loss is clearly warranted, so that valid and reliable assessments of audiologic performance can be obtained without the confounding factor of knowledge of English vocabulary.

Summary and Conclusions

This study assessed the effects of listener native language and age on recognition of unaccented English words and Spanish-accented English words. Monosyllabic words were presented in quiet and noise conditions. The principal findings were as follows:

1. L2 listeners consistently performed less accurately than L1 listeners on all English monosyllabic word recognition tasks used in the study. In addition, L2 listeners did not derive benefit from shared native language experience with the accented talker for recognition of Spanish-accented words. This finding was contrary to the prediction based on the ISIB for L2 listeners (ISIB-L). However, L2 listeners tended to show more modest declines in recognition of accented speech (re: unaccented English speech) compared to L1 listeners. This suggests that, although a talker's accent has a substantial impact on speech recognition by L1 listeners, the effect of talker accent is reduced for L2 listeners. Thus, native language experience in a language other than English has a large negative impact on recognition of unaccented and accented English words.
2. The effects of listener age were significant for L2 listeners, but not for L1 listeners. Older normal hearing L2 listeners performed less accurately than all other listener groups, indicating that these listeners have exceptional difficulty in recognizing English words in quiet and noise.
3. Cognitive performance on measures of working memory, speed of information processing, and inhibition did not predict monosyllabic word recognition performance when a unified linear mixed-effects regression model that controlled for random subject effects was applied to the data. However, the linguistic variable of English receptive vocabulary added significantly to the final model predicting listener performance. These results suggest that knowledge of the English lexicon is the most important variable contributing to performance by L1 and L2 listeners across the monosyllabic word conditions in quiet and noise tested in this study.

Taken together, the findings indicate that L2 listeners have more difficulty than L1 listeners understanding unaccented and Spanish-accented spoken English and that all listeners have more difficulty in understanding accented English compared to unaccented English. The results also show that older L2 listeners have considerable difficulty in all listening conditions. That is, older L2 listeners have quite poor recognition of unaccented English and Spanish-accented English monosyllabic words in both quiet and noise. These patterns of performance indicate that the native language of the listener, the native language of the talker, and the listener's age are key factors to be considered in understanding the nature of a listener's communication complaint. Familiarity with the English lexicon appears to play a key role in these findings, especially for L2 listeners. The results of this investigation should provide guidance for clinicians when evaluating older nonnative speakers of English in the clinical setting.

Acknowledgments

This research was supported by National Institute on Aging Grant R01AG009191, awarded to Sandra Gordon-Salant. The authors are grateful to Kelsey Oppler for her assistance in data collection and Stefanie Kuchinsky for her advice on the linear mixed-effects modeling analyses.

References

Adams, C., & Munro, R. R. (1978). In search of the acoustic correlates of stress: Fundamental frequency, amplitude, and duration in the connected utterance of some native and non-native speakers of English. *Phonetica, 35,* 125–156.

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-

impaired adults. *International Journal of Audiology, 47*(Suppl. 2), S53–S71. https://doi.org/10.1080/14992020802301142

American National Standards Institute. (2010). *S.36 Standard specification for audiometers* (Revision of ANSI S.36-1996, 2004). New York, NY: Author.

Anderson, S., White-Schwoch, T., Parbery-Clark, A., & Kraus, N. (2013). A dynamic auditory–cognitive system supports speech-in-noise perception in older adults. *Hearing Research, 300,* 18–32. https://doi.org/10.1016/j.heares.2013.03.006

Banks, B., Gowen, E., Munro, K. J., & Adank, P. (2015). Cognitive predictors of perceptual adaptation to accented speech. *The Journal of the Acoustical Society of America, 137,* 2015–2024. https://doi.org/10.1121/1.4916265

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mäechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1–48. https://doi.org/10.18637/jss.v067.i01

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America, 114,* 1600–1610. https://doi.org/10.1121/1.1603234

Bieber, R. E., & Gordon-Salant, S. (2017). Adaptation to novel foreign-accented speech and retention of benefit following training: Influence of aging and hearing loss. *The Journal of the Acoustical Society of America, 141,* 2800. https://doi.org/10.1121/1.4980063

Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of the Acoustical Society of America, 121,* 2339–2349. https://doi.org/10.1121/1.2642103

Broersma, M., & Scharenborg, O. (2010). Native and non-native listeners' perception of English consonants in different types of noise. *Speech Communication, 52,* 980–995. https://doi.org/10.1016/j.specom.2010.08.010

Burda, A. N., Scherz, J. A., Hageman, C. F., & Edwards, H. T. (2003). Age and understanding speakers with Spanish or Taiwanese accents. *Perceptual and Motor Skills, 97,* 11–20.

Calandruccio, L., & Smiljanic, R. (2012). New sentence recognition materials developed using a basic non-native English lexicon. *Journal of Speech, Language, and Hearing Research, 55,* 1342–1355. https://doi.org/10.1044/1092-4388(2012/11-0260)

Chiang, H.-H. (2018). English vocabulary size as a predictor of TOEIC listening and reading achievement among EFL students in Taiwan. *Theory and Practice in Language Studies, 8,* 203–212. https://doi.org/10.17507/tpls.0802.04

Clopper, C. G., Pisoni, D. B., & Tierney, A. T. (2006). Effects of open-set and closed-set task demands on spoken word recognition. *Journal of the American Academy of Audiology, 17,* 331–349. https://doi.org/10.3766/jaaa.17.5.4

Dubno, J. R., Dirks, D. D., & Morgan, D. E. (1984). Effects of age and mild hearing loss on speech recognition. *The Journal of the Acoustical Society of America, 76,* 87–96.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon identification of a target letter in a non-search task. *Perception & Psychophysics, 16,* 143–149.

Finch, W. H., Bolin, J. E., & Kelley, K. (2014). *Multilevel modeling using R*. Boca Raton, FL: CRC Press.

Flege, J. E. (2002). Interactions between the native and second-language phonetic systems. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 217–243). Trier, Germany: Wissenschaftlicher Verlag.

Flege, J. E., & Eefting, W. (1988). Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *The Journal of the Acoustical Society of America, 83,* 729–740.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Fox, R. A., Flege, J. E., & Munro, M. J. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *The Journal of the Acoustical Society of America, 97,* 2540–2551.

Füllgrabe, C., Moore, B. C. J., & Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition. *Frontiers in Aging Neuroscience, 6,* 347. https://doi.org/10.3389/fnagi.2014.00347

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge, United Kingdom: Cambridge University Press.

Gordon-Salant, S., & Cole, S. S. (2016). Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing. *Ear and Hearing, 37,* 593–602. https://doi.org/10.1097/AUD.0000000000000316

Gordon-Salant, S., & Friedman, S. H. (2011). Recognition of rapid speech by blind and sighted older adults. *Journal of Speech, Language, and Hearing Research, 54,* 622–631. https://doi.org/10.1044/1092-4388(2010/10-0052)

Gordon-Salant, S., Yeni-Komshian, G. H., & Fitzgibbons, P. J. (2010a). Recognition of accented English in quiet by younger normal-hearing listeners and older listeners with normal-hearing and hearing loss. *The Journal of the Acoustical Society of America, 128,* 444–455. https://doi.org/10.1121/1.3397409

Gordon-Salant, S., Yeni-Komshian, G. H., & Fitzgibbons, P. J. (2010b). Recognition of accented English in quiet and noise by younger and older listeners. *The Journal of the Acoustical Society of America, 128,* 3152–3160. https://doi.org/10.1121/1.3495940

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., Cohen, J. I., & Waldroup, C. (2013). Recognition of accented and unaccented speech in different maskers by younger and older listeners. *The Journal of the Acoustical Society of America, 134,* 618–627. https://doi.org/10.1121/1.4807817

Hale, S., Rose, N. S., Myerson, J., Strube, M. J., Sommers, M., Tye-Murray, N., & Spehar, B. (2011). The structure of working memory abilities across the adult life span. *Psychology and Aging, 26,* 92–110. https://doi.org/10.1037/a0021483

Hargus Ferguson, S., Jongman, A., Sereno, J. A., & Keum, K. A. (2010). Intelligibility of foreign-accented speech for older adults with and without hearing loss. *Journal of the American Academy of Audiology, 21,* 153–162. https://doi.org/10.3766/jaaa.21.3.3

Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics, 36,* 664–679. https://doi.org/10.1016/j.wocn.2008.04.002

Imai, S., Walley, A. C., & Flege, J. E. (2005). Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. *The Journal of the Acoustical Society of America, 117,* 896–907. https://doi.org/10.1121/1.1823291

Janse, E. (2009). Processing of fast speech by elderly listeners. *The Journal of the Acoustical Society of America, 125,* 2361–2373. https://doi.org/10.1121/1.3082117

Janse, E., & Adank, P. (2012). Predicting foreign-accented adaptation in older adults. *Quarterly Journal of Experimental Psychology, 65,* 1563–1585. https://doi.org/10.1080/17470218.2012.658822

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19,* 1–36.

Martin, N. A., & Brownell, R. (2011). *Receptive One-Word Picture Vocabulary Test–Fourth Edition*. Novato, CA: ATP Publications.

Mayo, L. H., Florentine, M., & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language, and Hearing Research, 40,* 686–693.

Munro, M. J. (1998). The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition, 20,* 139–154.

Nakamura, K., & Gordon-Salant, S. (2011). Speech perception in quiet and noise using the hearing in noise test and the Japanese hearing in noise test by Japanese listeners. *Ear and Hearing, 32,* 121–131.

Parbery-Clark, A., Strait, D. L., Anderson, S., Hittner, E., & Kraus, N. (2011). Musical experience and the aging auditory system: Implications for cognitive abilities and hearing speech in noise. *PLoS One, 6,* 1–8. https://doi.org/10.1371/journal.pone.0018082

Pinet, M., & Iverson, P. (2010). Talker–listener accent interactions in speech-in-noise recognition: Effects of prosodic manipulation as a function of language experience. *The Journal of the Acoustical Society of America, 128,* 1357–1365. https://doi.org/10.1121/1.3466857

Ramkissoon, I., Proctor, A., Lansing, C. R., & Bilger, R. C. (2002). Digit speech-recognition thresholds (SRT) for non-native speakers of English. *American Journal of Audiology, 11,* 23–28.

Schurman, J., Brungart, D. S., & Gordon-Salant, S. (2014). Effects of masker type, sentence context, and listener age on speech recognition performance in 1-back listening tasks. *The Journal of the Acoustical Society of America, 136,* 3337–3349. https://doi.org/10.1121/1.4901708

Shi, L.-F. (2010). Perception of acoustically degraded sentences in bilingual listeners who differ in age of English acquisition. *Journal of Speech, Language, and Hearing Research, 53,* 821–835. https://doi.org/10.1044/1092-4388(2010/09-0081)

Stuart, A., & Phillips, D. P. (1996). Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbyacusic listeners. *Ear and Hearing, 17,* 478–489.

Tabri, D., Abou Chacra, K. M., & Pring, T. (2011). Speech perception in noise by monolingual, bilingual, and trilingual listeners. *International Journal of Language & Communication Disorders, 46,* 411–422. https://doi.org/10.3109/13682822.2010.519372

Tillman, T. W., & Carhart, R. (1966). *An expanded test for speech discrimination utilizing CNC monosyllabic words: Northwestern University Auditory Test No. 6*. USAF School of Aerospace Medicine, Report No. SAM-TR-66-55.

Vaughan, N., Storzbach, D., & Furukawa, I. (2006). Sequencing versus non-sequencing working memory in understanding of rapid speech by older listeners. *Journal of the American Academy of Audiology, 17,* 506–518.

von Hapsburg, D., Champlin, C. A., & Shetty, S. R. (2004). Reception thresholds for sentences in bilingual (Spanish/English) and monolingual (English) listeners. *Journal of the American Academy of Audiology, 15,* 88–98.

Weber, A., Broersma, M., & Aoyagi, M. (2011). Spoken word recognition in foreign-accented speech by L2 listeners. *Journal of Phonetics, 39,* 479–491. https://doi.org/10.1016/j.wocn.2010.12.004

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale–Third Edition (WAIS-III)*. San Antonio, TX: The Psychological Corporation.

Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., . . . Gershon, R. C. (2013). Cognition assessment using the NIH toolbox. *Neurology, 80,* S54–S64. https://doi.org/10.1212/WNL.0b013e3182872ded

Wenk, B. J. (1985). Speech rhythms in second language acquisition. *Language and Speech, 28,* 157–175.

Xie, X., & Fowler, C. A. (2013). Listening with a foreign-accent: The interlanguage speech intelligibility benefit in Mandarin speakers of English. *Journal of Phonetics, 41,* 369–378. https://doi.org/10.1016/j.wocn.2013.06.003

Yeni-Komshian, G. H., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism: Language and Cognition, 3,* 131–149.